DOCUMENT RESUME

ED 217 796 HE 015 198

AUTHOR Moran, Tom; Rowse, Glenwood
TITLE Quantitative Procedures for the Assessment of Quality in Higher Education Institutions.
INSTITUTION New York State Education Dept., Albany. Office of Postsecondary Research, Information Systems, and Institutional Aid.
SPONS AGENCY Ford Foundation, New York, N.Y.
PUB DATE Dec 81
GRANT FF-805-0540
NOTE 63p.; Some tables may not reproduce well due to light or small print.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Ancillary School Services; College Role; *Comparative Analysis; *Educational Assessment; *Educational Quality; *Evaluation Methods; *Higher Education; Institutional Characteristics; Outcomes of Education; Research Projects; Statistical Analysis; Student Development; *Systems Approach
IDENTIFIERS New York

ABSTRACT

The development of procedures designed to provide quantitative assessments of quality in higher education institutions are reviewed. These procedures employ a systems framework and utilize quantitative data to compare institutions or programs of similar types with one another. Three major elements essential in the development of models focusing on the assessment are as follows: they must be multidimensional, they must consider quality in relationship to the purposes and mission of the institution, and comparisons should be made of institutions within comparable groups; and judgments about quality intended for use in policy considerations should rest on quantifiable data tied as closely as possible to firm outcome measures. The approach was developed to assess quality within three generally accepted primary missions of higher education: student development, research, and service. A matrix may be used to display relevant data (i.e., categories for major institutional purposes on the horizontal axis and basic institutional variables or descriptors on the vertical axis). These descriptors serve as outcome indicators for the three institutional purposes areas. The systems elements of outcomes, inputs, and transformation process are included, and the elements for each category are specified. The data enable longitudinal self-comparison, snapshot comparison with normative cutoffs or ranges, and comparisons with groups of other similar institutions. The procedures are elaborated and an example of the deviation from the mean approach is presented, based on faculty and tuition data from all New York state institutions between 1977-1979. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED217796

QUANTITATIVE PROCEDURES FOR THE ASSESSMENT OF QUALITY IN HIGHER EDUCATION INSTITUTIONS

DECEMBER 1981

This report summarizes the development of procedures designed to provide quantitative assessments of quality in institutions of higher education. It was prepared under a planning grant from the Ford Foundation (grant #805-0540). This report does not necessarily reflect official positions of the Ford Foundation, the New York State Education Department, or the Board of Regents.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Paul Weinj

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

HE 015 196

Office of Postsecondary Research, Information Systems and Institutional Aid
New York State Education Department
Cultural Education Center, Room 5B44
Albany, NY 12230

PREFACE

The assessment of institutional quality is one of the major policy considerations that higher education must confront in the 1980's. An era of dramatic growth and expansion has come to an end and will likely be succeeded by a period of significant contraction for the system of postsecondary education. The decline in the size of the system as a whole and the concommittant reduction of education as a national priority will result in increasing scarcity of resources. The conditions of scarcity dictate that not all educational endeavors can be funded and therefore that not all such endeavors can endure. A higher education community serving this society's interests must insure that its highest quality institutions and programs survive.

But how is quality to be defined; and by whom; and once defined, how is it to be assessed? These questions are the foci on which policy decisions must center. In an effort to begin to address these questions, the New York State Education Department with partial support from a Ford Foundation grant initiated research in this area.

This research resulted in the development of procedures for assessing quality which are described in this paper. These procedures employ a basic systems approach and utilize quantitative data to compare institutions of homogenous types with one another. It should be stressed that the approach makes no pretense to measure quality in absolute terms and is only a device for comparing the manner in which institutions deviate from the norm. It should also be stressed that the approach is highly tentative and incomplete. It will require extensive research and development before its accuracy can be validated and its utility evaluated.

The approach to the assessment of "quality" outlined in this paper was developed by a research team working under the supervision of Paul Wing, Coordinator of the Office of Postsecondary Research and Information Systems and Institutional Aid. This paper was prepared by Tom Moran and Glenwood Rowse. Questions can be addressed to Paul Wing at (518) 474-5091.

INTRODUCTION

The significance of quality in American Higher Education is unparalleled. As Thomas Stauffer explains in the introduction to Quality: Higher Education's Principal Challenge which was recently published by the American Council on Education,"public confidence is the most precious asset that the higher education community possesses." That confidence, however, can only be maintained if quality in the academic and other missions entrusted to these institutions is continued. The message in Stauffer's observation is straightforward and unambiguous: quality is indisputably important in higher education.

Despite general agreement on the centrality of quality, there is no general agreement on how it should be defined or assessed. This report describes a method for assessing quality which relies upon quantifiable data organized in a systems matrix.

Methods of Assessing Quality

Traditionally there have been four major methods for assessing quality. The most well known of these methods are the reputational studies which prepare rankings of the nation's major graduate programs, and recently, undergraduate programs. The obvious weakness in this method is that it uses no standards of quality which are explicit. Each evaluator renders his/her overall judgement of a program based on whatever criteria he/she feels are important. A further weakness in this approach is that it incorporates evaluations of fewer than 1% of the schools in the country.

Perhaps the most common method of quality assessment is the type undertaken by the various regional accreditation agencies throughout the country. These evaluations generally amount to self-study procedures undertaken by the institution in accordance with guidelines established by the accrediting agency. These evaluations generally focus more on the resources and the processes operating within the institution than on other issues. Rarely, however, do these studies address issues related to the quality of outcomes.

Another well known method of assigning quality to institutions is to examine composite profiles of their student input characteristics. The assumption here is that high quality students both reflect and help to create high quality institutions.

The most comprehensive orientation to evaluating quality is found in the Gourman Report (1977). This method utilizes a wide range of factors in assessing quality but doesn't stipulate the criteria by which judgements on these factors are made. Scores given to each factor are therefore highly dependent on the subjective valuations of the individual making the judgements. This lack of inter-rater reliability is the fundamental weakness of this method. The numerical rankings of institutions obtained by this method are merely an illusion of empirical soundness, but they lack rigor and validity.

Each of these methods contains major weaknesses. Either they are uni-dimensional and do not incorporate a large enough range of variables to give a reliable judgement on quality, or they are reasonably comprehensive but subjective in character. In the latter case, the judgements rest on more complete observations but use criteria that are not made explicit and, therefore, cannot be reviewed, much less duplicated, by other researchers. In almost all past efforts, the direct measurement of student learning has been underattended in favor of more easily approached outcomes such as prestige and reputation or research grants and publications.

Essential Elements of Quality Assessment

As Donald Hester (1980) makes clear, at least three elements are essential in any paradigm established for determining quality 1980. The first is that quality is a multi-dimensional issue; no single factor can be used to assess it. For instance, judgements about institutional quality that rest entirely on information about student input characteristics are inevitably incomplete. It may, of course, be true that student characteristics reveal a high positive correlation with institutional quality and therefore can serve as useful proxy measures for quality. However, in order to establish this correlation, it is necessary to possess a comprehensive and reliable assessment of quality for which student characteristics can serve as a proxy.

Of course, the task for the researcher, which is both basic and monumental, is to delineate the model which reveals "a comprehensive and reliable assessment of quality." Once this task is completed, any single measure or cluster of measures which indicate a high correlation with "quality" can then be used as proxies.

The second element which is necessary in models which assess the quality of institutions is that the institutions must be evaluated in accordance with their objectives or missions. In other words, a community college with no emphasis on research cannot be evaluated by the same standards used to evaluate a major research university. The traditional emphasis on a single status system for determining quality is at the core of the reputational studies discussed earlier. This system contains the implicit assumption that all institutions should be evaluated against the standard of the elite research-oriented university. It is, as has been stated, an inappropriate paradigm for broad scale assessment of institutional quality that includes reviews of a great many institutions of varying sizes and types. For a valid effort, it is necessary to combine a system for evaluating quality with a classification scheme that considers quality in relationship to groupings of institutions by common characteristics. The common characteristics can be fairly arbitrary but should give considerable weight to the factors of size and mission.

Finally, quality assessment must rest on some empirical standards. This is not to deny the ultimate importance of philosophical propositions relating to issues of value, but once these have been established, it is necessary to develop a system of evaluation that is supported by reliable quantifiable data whether objective or judgemental in nature. It must be understood that we are not discussing the type of decisions that individuals might make in answering the highly personal questions of whether or not they <u>like</u> a particular institution. Very possibly, quantifiable data used to answer that question, no matter how copious, would be irrelevant to

assessment of quality. Rather, we are attempting to develop a methodology by which planners and administrators both at the institutional and regional level can make policy decisions related to issues of quality. For this purpose, it is necessary to obtain data that can be used for purposes of institutional comparison. Moreover, the data set used in these situations must be based on its suitability for assessing quality not simply on its availability. In other words, a model for assessing quality, while it cannot be developed independently of questions related to data availability, cannot be designed solely around available data either.

In summary, there are three major elements essential in the development of models focusing on the assessment of quality: First, they must be multidimensional, second, they must consider quality in relationship to the purposes and mission of the institution and comparisons should be made of institutions within comparable groups; third, judgements about quality intended for use in policy considerations should rest squarely on quantifiable data that are tied as closely as possible to firm outcome measures. The approach developed in this project meets all of these considerations.

THE QUALITY INDICATORS APPROACH

The approach described below was developed using basic systems concepts to assess quality within three generally accepted primary missions of higher education.

1) Student Development,

2) Research, and

3) Service.

The matrix in which data may be displayed (see Figure 1 on the following page) contains categories for major institutional purposes on the horizontal axis and a variety of basic institutional variables or descriptors on the vertical axis. These descriptors serve as outcome indicators for the three institutional purposes areas. These variables are arranged within a systems analysis framework. The familiar systems elements of Outcomes, Inputs, and Transformation Process are all included and the elements for each category are specified. Variables under the outcome heading reflect direct measures while the input and process variables are indirect measures of the outcomes. We have used the term 'descriptors' for these variables because, until data are collected and interpreted, they are no more than general categories to describe the institution. The listing of a few of the descriptors within the outcome category shown in Table 1 should illustrate. For a complete explanation of all the descriptors and their arrangement in the matrix, see Appendix A.

As can be seen, the descriptors are simply the element or aspect of an institution under scrutiny. However, when data are obtained for each descriptor and a range or cut-off value is established for it, the descriptor becomes an indicator. For instance, if the descriptor "value added" is examined for each institution within a particular institutional type and a cutoff value is established below which an institution can be said to be not achieving sufficient results, then the "value added" descriptor becomes an "indicator" of one aspect of quality.

Institutional Type:

Figure 1

Program or Institutional Outcome Area

MEASURE & INDICATOR AREAS	LEARNING: Student Development (KAS) for Each Curric Area or: Life	Career	Specalized	RESEARCH: Knowledge Development by Program: Basic	Applied	SERVICE: Public Benefit: Comm.	Bus/Ind	Society	Critica: Inst. Value	Cutoff Level
OUTCOME										
Learning										
Outcome Level										
Value Added										
Placement										
Completion										
Social Behavior										
Satisfaction										
R & D										
Products										
Grants										
Service										
Community Projects										
Consulting										
Productions										
INPUT (Student Targets)										
Ability										
Motivation										
Need										
ENVIRONMENT/PROCESS										
General Indicators										
Attractiveness										
Reputation										
Effort										
Instruction										
Student/Method Match										
Application of Model Selected										
Teacher-Student Interaction										
Time Spent										
Curriculum										
Organizational Health										
Adaptability										
Morale										
Nature of Decision-making										
Nature of Planning/Mgt.										
Communication										
Incentive Structure										
Preservation of Expertise/ecc.										
Allocation of Resources to Improvement/Maintenance										
Resources										
Faculty										
Critical Mass for Planning										
Flexicility										
Vigor										
Financial										
Allocation to Goals										
Level to Goals										
Trends										
Facilities										
Amount										
Condition										
Trends										
Admin/Support Staff										
Amount										
Nature										
Trends										
Support Structures										
Counseling										
Admissions										
Remediation										
Financial Aid										
Services										
Student Center										
Day Care Center										
Health Care										
Transportation										
Climate/Atmosphere										
Emphasis on Scholarship Proprietary, etc.										
Intimacy										
Moral/Ethical										
Complex Variables										
Summated Triggers										
Coincident Conditions										

One Matrix for Plans of Inst. or Our Criteria and One Matrix for Actual Results

Table 1. Possible Outcome Measures to Use in Higher Education Assessment

OUTCOME CATEGORY	POSSIBLE MEASURES
Outcome level	Comparison with national norms on such tests as GRE's
Value added	Gain scores on standardized test, etc.
Placement	Employment profiles of graduates
Completion rate	Retention figures
Social behavior	Participation in Public Affairs, etc.
Satisfaction	Student satisfaction
Research and Development	
Products	Patents, citations, awards, publications, etc.
Grants	Number and dollar amounts of grants
Services	
Community projects	Number and impact of projects
Consulting of faculty	Number of consultancies
Student Targets	
Ability	As measured on standard indicators
Motivation	As measured by self-assessment
Need	Financial need

If the matrix in Appendix A is examined, it will be seen that a Utilization category appears on the horizontal axis as well as the Purposes category mentioned earlier. Within the utilization category there are sub-categories identifying the group(s) for whom this particular indicator will be useful, including state agencies, regional planning groups, campuses, and academic review units. Each of these groups has quite different purposes and intentions with respect to data on quality. Campuses, for instance, are interested primarily in formative evaluation that can assist them in improving their performance. Academic review units, on the other hand, are interested primarily in summative evaluation that enables them to determine whether an institution or program has met a basic threshold quality.

From an institutional point of view, data collected according to this framework enables three kinds of comparisons: 1) historical/longitudinal self-comparison on the basis of trend data; 2) "snapshot" comparison with normative cutoffs or ranges; and 3) comparisons with groups of other similar institutions.

All of the indicators used in the initial development of this project were drawn from existing data bases available through the Office of Postsecondary Research and Information Systems in the New York State Education Department. Consequently, it was possible to manipulate the data and provide quantitative analysis that yielded profiles of how institutions related to other institutions of similar type and mission. These profiles utilize standardized scores (z) and can be devised at the indicator level, or in aggregate form to compare institutions in their total profile either with norms for their type or with all other institutions. A z score (or standard score as it is sometimes called) identifies the position of the score relative to the mean of the distribution using the standard deviation as the unit of measurement. Use of the z score enables researchers working with this system to assess an institution's position, with respect to a given descriptor, relative to other institutions. We must caution again that we are not attempting measuring quality in any absolute sense, but rather simply comparing the differences between institutions. The relationship

quality and indirect indicators will be an important research agenda in any future research following up on this project.

Conceptual Foundations of the Model

The following principles and issues served as guidelines and conceptual foundations in the development of the model on which this quality assessment matrix is based:

1) Although individual descriptors of an institution may be unreliable indicators of the institution's status or performance, by using a large number of appropriate descriptors it is expected that a relatively reliable picture of the overall status can be constructed.

2) When screening institutions or programs (for excellence, weakness, etc.), the cutoffs or acceptable ranges applied to specific measures or indicators should be determined separately for different types of institutions.

3) A single indicator may be useful in assessing more than one purpose, but a distinct cutoff level must be established for each purpose about which the indicator provides information.

4) Research regarding measures and indicators appropriate for a given mission/task may be undertaken independent of institutional type or peer group. Only the cutoff levels used when screening institutions are dependent on institutional type.

5) Institutional assessment and program assessment are distinctly different problems requiring different sets of indicators. The diverse nature of individual institutions may seriously limit the utility of aggregate institutional assessments. However, it seems inevitable that institutional assessments will continue to be made.

6) Because of the varying constellation of mission areas, the number of significantly different institutional types is almost certainly large (minumum 30 and more likely up to 200 in both cases) if one wishes to obtain a reasonable degree of homogeneity within the types. If the procedures described in this report are used by a State agency to measure the extent to which institutions achieve desirable results, group comparisons (distributions) may have limited utility as a means for assessing such desireability because the groups may be too small. In any case, desirability on a given measure should be determined with the participation and cooperation of the institutions themselves.

7) Indicators appropriate for campuses (in terms of utility, cost, and feasibility) for self-assessment and improvement will be different from those useful to state agencies undertaking measurements against minimum standards. The institutions will be more interested in process and excellence, while state agencies may focus on minimum standards as a general screen and regard excellence as an issue for further analysis.

8) Alternative uses for the indicators should be identified and developed to increase their utility and to obtain a cost/benefit ratio that will support their development. Examples of uses of quality measures and indicators include:

   a) Self-assessment and improvement by individual institutions.

   b) Public distribution of institutional and aggregated results as a means of

      a) Providing information to the public

      b) Improving student consumer decisions

      c) Encouraging less disruptive and less costly competition among institutions

      d) Regulation and State policy decisions.

9) One should work from a comprehensive conceptual framework when assessing quality or identifying indicators. To do otherwise is to continue the risk of ignoring important outcome areas (such as student learning).

10) Campus representatives would be invaluable in identifying undesirable side effects, disincentives, or misinterpretations that would accompany the use of any criteria proposed for use in assessing program or institutional excellence.

11) Consensus from all user groups on purpose, method, and results should be sought as part of any project building on the ideas in this report. This consensus should be related to the usability of initial results and the desirability of continued development, rather than on the perfection of the results.

IMPLEMENTATION PROCEDURES

This section of the report describes specific strategies and procedures for implementing the concepts outlined above. A complete development would involve tens, perhaps hundreds, of specific measures each linked to general variables and carefully considered in the context of the missions and goals of specific institutions. Because of limitations on the resources of this project, it has been possible to develop only one basic dimension of the overall conception, specifically faculty resources. This does, however, illustrate the approach and provide a basis for considering its advantages and disadvantages as an assessment procedure.

Once the details of the faculty resources example have been developed, suggestions are provided for developing aggregate, summary profiles that offer some promise as the basis for a variety of assessment and planning activities, at both the campus and state levels.

To add realism and impact to the presentation, real data for colleges and universities in New York State have been used in the examples that follow. Readers should recognize that this has been done to emphasize the technical feasibility of the procedures. It does not mean that any discussions or agreements have been reached about the appropriateness of the specific measures or utility scales chosen. Nor does it imply that the measures chosen represent the full set of relevent measures. There are undoubtedly many other measures that would be relevant and desireable, but in the absence of appropriate data they have had to be omitted from the analyses, at least at this time.

Elaboration of the Procedures and An Example of the Deviation from the Mean Approach

Faculty and tuition data were run through a complete set of analyses to clarify and illustrate the Deviation from the Mean approach and explore its utility. A Data set was available for all New York State institutions between 1977 and 1979. Five faculty measures were included of which the trend measure was available for four.

The measures included were: students/faculty, mean salary, percent full-time faculty, percent tenured faculty, and faculty per program. Undergraduate and graduate tuition and fees were included as well. Definitions for each variable are provided in Appendix A to assist those interested in assessing the relevance of the operational measures used in the study.

The use of Z scores may be accomplished in several ways. A purely descriptive procedure would involve simply submitting the raw scores to the analysis. A value laden procedure would involve some conversion of the raw scores to reflect their relationship to the specific interest at the moment, whether it be overall quality, instructional quality, research quality, or some other outcome. The simplest procedure is to assign a utility measure to each variable. This involves no more than a specification of positive or negative value indicating that more or less of a variable is good. For example, a low student-faculty ratio is generally considered to be more conducive to high instructional quality than is a high ratio. Therefore, when examining this indicator, each ratio would be multiplied by minus one before entering the data for analysis (or the sign may be specified as part of the analysis). Even with this relatively unsophisticated approach to assigning value, experience humbles us. It is often difficult to obtain agreement on the appropriate sign to use, it is not always clear what is desirable and a utility measure can have a different value depending on what outcome is being examined. For example, high faculty salaries may be very positive when examined as a quality indicator in the instructional category. In this instance, high salaries may be construed as a comparative advantage for an institution because it reflects the institution's capacity to attract teachers of high quality. On the other hand, when examined as a measure in the financial status category, high salaries may be a negative attribute because the institution's salaries may create financial strains. One thing does seem clear: the more specific the goal and situation for which the measure will be used, the easier it is to agree on a sign. For general purposes involving multiple goals and situations, the descriptive procedure is more flexible and appropriate.

A more rigorous value laden procedure would attempt to transfer the raw data scale to an ordinal quality scale, preferably with uniform intervals. Detailed and laborious resource would be necessary to begin even identifying interval scales. However, useful approximations might be obtained by tapping expert judgment. One must consider, however, that the overall analysis may tend to fragment, since it will be driven toward very specific outcome goals by this procedure.

Variables:

Provided below are illustrative descriptions for the variables included in this report which emphasizes faculty resources related to learning outcomes. A utility measure reflecting a simple direction of goodness was assigned to each variable with respect to its overall influence on student learning or instructional quality; and the rationale upon which the value scale was determined for each variable is summarized to suggest the kind of discussions and negotiations that might go on. Before this general approach was implemented in practice, these descriptions would be refined and supported by a variety of studies and discussion sessions. These should elucidate the logic governing any tradeoffs that may be necessary when establishing priorities among a set of goals or outcome areas, and also make clear any shortcomings of the measures.

Students Per Faculty: With respect to learning outcomes this measure is a proxy for the personal attention received by individual students. Increased attention is usually thought to both improve motivation and facilitate the mechanics of learning. Lower ability and less internally motivated students may have more need for this kind of attention than bright, internally motivated students. Students per faculty measures a necessary but insufficient condition for both in-class and out-of-class attention, but it is not clear how highly it correlates with either.

Considerable research has been done on class size. It appears that class size may vary between approximately 12 and 35 without much difference in the level of personal attention received. Class sizes either above or below that range seem to have different dynamics. Given that both cost of instruction and research productivity of faculty might be hindered by increased student contact with faculty, there is a tradeoff to be considered when setting goals for this variable. For these analyses which focus on student learning, a low ratio is considered desirable and a minus sign has been assigned to the variable.

Salary: Ten month salaries were adjusted to provide 12 month annual average salaries for an institution to obtain a measure comparable across institutions. The measure still has problems in terms of comparability, however, due to the different emphases that colleges place on full- and part-time faculty and due to differences in their tenure and rank profiles. These differences by themselves can produce quite different average salaries for different institutions. For the purpose of these analyses, however, the comparability was considered adequate.

With respect to learning outcomes, average salary is thought to be related to the quality of faculty members which can be attracted to the institution. In turn and on the average better curriculum and instruction should be present where better salaries are paid. It may be particularly important to make comparisons on the variables within general types of institutions. The direction of goodness assigned to this variable was a plus sign.

Percent Tenured: The proportion of faculty who are tenured is thought to be related to learning in a couple of ways. We are living in a knowledge explosion at present. It has frequently been suggested that colleges need a regular influx of newly trained people and fresh ideas to maintain quality programs. While literature does not provide firm conclusions on this issue, it is an issue deserving consideration. Percent tenured is

a useful proxy for an institution's ability to provide a steady influx of new ideas. It is an indication of flexibility in another way as well. A highly tenured institution may have difficulty adapting its programs as the interests of its clientele change and unless it is rich it may be forced into limiting certain student services and instructional costs in order to pay the high average salaries of its faculty. While some of these problems are small during stable times, major program shifts are underway in colleges now.

A low faculty percent tenured may possibly indicate a willingness to hire whoever is available (lower average quality) since the security of tenure is an attraction for even the best faculty members. Thus, the ideal average level of tenure for an institution may be a moderate to low level. For this analysis, a negative sign was assumed as the appropriate direction of goodness with no test applied to identify those who might be too low.

Percent Full-Time: Conventional wisdom suggests that full-time faculty members contribute more to departmental planning and operational tasks and give more time outside of class to students. In turn, there should exist stronger programs and greater learning. The literature seems to provide little evidence on this matter.

Issues of critical mass may also be important but there may be a variety of ways of obtaining a critical mass for a department or institution. In general, however, an institution with a high percentage of full-time faculty should be in a strong position. Consequently, a plus sign was assigned to this measure.

Faculty per Program: Full-time equivalent faculty divided by the number of four-digit HEGIS programs was calculated to get at critical mass for providing programs. Too low a ratio might indicate that faculty are stretching their expertise and resources too thin across the programs of the institution. Of course, effective coordination of courses to serve

students across program areas can possibly lower the ratio needed for critical mass. The extensive use of either adjunct faculty or graduate assistants may be either a positive or a negative factor depending on a variety of conditions. Alterations of the exact faculty per program ratio measured might be considered due to these kinds of contingency influences. Overall, a high faculty per program ratio may be considered a sign of, or at least a precondition for strength. The direction of goodness assigned to this variable was a plus sign.

Tuition & Fees: Undergraduate and graduate tuition and required fees were separately included in the analyses. While differences on these measures between institutions from different sectors or types of institutions are clearly not directly useful as indicators of learning, differences within type of institution may be meaningful. The stronger institutions with the better reputations should be able to charge higher tuition levels. It might be desirable to generate an interaction variable which places high value on high tuition only when the acceptance ratio is low. For these analyses a plus sign was used, but tuition and fees were not included in any of the average or total measures developed.

Summary Analysis of the Data:

Whether value scales are applied to the raw data or not, it is useful to obtain means and ranges for each variable for each type of institution. This information provides reference points that are useful whether or not additional analyses are conducted. For these summary statistics to be truly useful, the institutional types should be internally homogenous, that is, consist of comparable institutions. Considerable work remains to be done in developing an appropriate institutional typology for this type of analysis. For this example, however, a typology recently developed in New York State was used. It classifies institutions according to the following categories: 1) Multiversities, 2) Universities,

3) College Complexes, 4) Colleges, 5) Specialized Colleges, 6) Seminaries, 7) Health and Professional Schools, 8) Nursing Schools, 9) Two Year Schools. The resulting types are reasonably homogenous except for the Specialized category. The specialized category is a mixture of specialized institutions which did not fit elsewhere and data for that category should be used with extreme caution.

Tables 2 to 4 on the following pages provide the summary statistics for the variables included in this example. Since both the current magnitude of a variable and the trend for the variables are considered important, both are included in the summary tables. For most of the variables, there exists considerable variation for both level and trend measures, both within and across types of institutions. If one assumes that the types are in fact homogeneous groups, the existence of large varations may be an indication that the variable is inappropriate, or that there are strong and weak institutions within the types, or simply that there are diverse ways of managing an institution of a given type. Further research, including validation by experts, will be needed to get a clearer determination for each variable.

Generally speaking, however, if excessive deviation from the norm is evident or if there is a significant change in trend data for a particular variable, then further scrutiny is warranted. For example, a 30 percent decline in average salary at a university (as shown in the table on the next page) is probably a clear signal of some sort of problem.

Summary data of this nature is rarely made available to institutional planners. If developed for a fairly comprehensive set of variables and viewed for a group of institutions, the composite "picture" provided by these statistics would enable planners to see more clearly the shape of the higher education system. This would also presumably permit many decisions to be made on a more rational basis.

TABLE 2 SUMMARY STATISTICS

TYPE OF INSTITUTION	AVERAGE SALARY OF FACULTY			STUDENT FACULTY RATIO			STUDENT FACULTY RATIO WITH TAS					
	MEAN	HIGH	LOW	MEAN	HIGH	LOW	MEAN	HIGH	LOW	N	N	N
1979 LEVEL*												
	$	$	$	#	#	#	#	#	#			
MULTIVERSITIES	30035.	34337.	27211.	12.	19.	4.	9.	14.	4.	8	8	8
UNIVERSITIES	28829.	32825.	23477.	17.	22.	13.	16.	20.	12.	18	18	18
COLLEGE COMPLEXES	25243.	32543.	19300.	16.	29.	10.	16.	29.	10.	41	41	41
COLLEGES	20310.	28858.	14278.	15.	27.	10.	15.	27.	10.	36	36	36
SPECIALIZED	25096.	46374.	2802.	16.	47.	3.	14.	29.	3.	16	16	16
SEMINARIES	18723.	35663.	3304.	7.	12.	3.	7.	12.	3.	11	18	18
HEALTH PROFESSIONAL	23640.	29633.	17118.	11.	51.	1.	11.	51.	1.	4	9	9
NURSING SCHOOLS	20112.	21598.	18199.	9.	13.	7.	9.	12.	7.	6	6	6
TWO-YEAR SCHOOLS	20256.	33730.	9779.	21.	44.	10.	21.	44.	9.	77	82	92
NEW YORK STATE	$22615.	$46374.	$ 2802.	17.#	51.#	1.#	16.#	51.#	1.#	217	234	234
1977-1979 % CHANGE**												
MULTIVERSITIES	13.	18.	8.	-3.	5.	-12.	-6.	19.	-20.	8	8	8
UNIVERSITIES	14.	22.	7.	-4.	28.	-29.	-5.	25.	-31.	18	18	18
COLLEGE COMPLEXES	15.	29.	6.	-3.	23.	-32.	-2.	23.	-33.	41	41	41
COLLEGES	12.	34.	-1.	-2.	31.	-35.	-2.	31.	-35.	36	36	36
SPECIALIZED	7.	32.	-68.	4.	232.	-41.	4.	232.	-41.	16	16	16
SEMINARIES	19.	48.	-11.	5.	239.	-65.	4.	239.	-65.	10	18	18
HEALTH PROFESSIONAL	11.	22.	2.	-9.	17.	-29.	-6.	19.	-29.	4	9	9
NURSING SCHOOLS	24.	47.	3.	-7.	49.	-30.	-7.	49.	-30.	4	4	4
TWO-YEAR SCHOOLS	12.	27.	-28.	-4.	51.	-91.	-4.	51.	-91.	73	80	90
NEW YORK STATE	13.%	48.%	-68.%	-2.%	239.%	-91.%	-3.%	239.%	-91.%	210	230	230

SOURCE: NYSED,POSTSECONDARY RESEARCH INFORMATION SYSTEMS 05/12/81

* 1978 DATA USED IF 1979 NOT AVAILABLE
** 1976-1978 DATA USED IF 1977-1979 NOT AVAILABLE

TABLE 3 SUMMARY STATISTICS

TYPE OF INSTITUTION	% TENURED FACULTY PER FT INSTR FACULTY			% FULL-TIME FACULTY EXCLUDES TAS			FTE FACULTY PER PROGRAM					
	MEAN	HIGH	LOW	MEAN	HIGH	LOW	MEAN	HIGH	LOW	N	N	N
1979 LEVELS*	%	%	%	%	%	%	#	#	#			
MULTIVERSITIES	56.7	69.0	38.0	79.1	92.4	45.5	5.3	9.9	1.7	3	3	3
UNIVERSITIES	66.3	86.0	46.0	64.3	90.3	33.8	4.5	7.6	2.2	13	13	13
COLLEGE COMPLEXES	61.6	94.0	17.0	65.7	99.2	20.5	3.5	12.2	1.2	41	41	41
COLLEGES	41.6	77.0	0.0	61.0	90.1	8.3	4.4	37.4	1.1	34	34	36
SPECIALIZED	24.0	84.0	0.0	43.5	98.3	1.6	9.7	55.0	0.7	17	16	16
SEMINARIES	11.2	75.0	0.0	52.0	96.8	14.9	5.3	9.0	2.8	19	17	18
HEALTH PROFESSIONAL	12.7	58.0	0.0	52.8	95.0	2.7	66.9	434.0	10.0	9	9	9
PROPRIETARY SCHOOLS	0.0	0.0	0.0	97.6	100.0	95.7	11.5	23.0	3.0	6	6	6
TWO-YEAR SCHOOLS	38.7	95.0	0.0	57.0	100.0	6.9	6.9	92.0	0.5	83	80	82
NEW YORK STATE	% 40.6	% 95.0	% 0.0	60.1%	100.0%	1.5%	8.2#	434.0#	0.5#	237	231	234
1977-1979 % CHANGE**												
MULTIVERSITIES	-1.5	5.7	-10.5	-2.0	10.8	-14.3				3	3	0
UNIVERSITIES	7.3	48.5	-17.1	-0.0	24.1	-15.5				13	13	0
COLLEGE COMPLEXES	1.8	41.0	-26.3	4.1	220.9	-24.4				41	41	0
COLLEGES	1.4	52.5	-40.3	5.4	60.1	-22.3				34	34	0
SPECIALIZED	0.2	44.2	-22.0	-1.9	109.9	-91.9				8	16	0
SEMINARIES	12.8	96.1	-13.5	-15.6	91.4	-55.9				6	17	0
HEALTH PROFESSIONAL	12.4	69.6	-49.4	9.9	143.9	-55.3				4	9	0
PROPRIETARY SCHOOLS			1000.0	9.9	50.0	-14.3				0	4	0
TWO-YEAR SCHOOLS	12.1	133.7	-60.0	2.4	153.9	-50.0				45	74	0
NEW YORK STATE	5.6%	133.7%	-60.0%	1.6%	220.9%	-91.9%				164	225	0

SOURCE: NYSED, POSTSECONDARY RESEARCH INFORMATION SYSTEMS 04/23/81

* 1978 DATA USED IF 1979 NOT AVAILABLE
** 1976-1978 DATA USED IF 1977-1979 NOT AVAILABLE

TABLE 4: SUMMARY STATISTICS

TYPE OF INSTITUTION	UNDERGRADUATE TUITION & FEES			FIRST-PROFESSIONAL TUITION & FEES			GRADUATE TUITION & FEES					
	MEAN	HIGH	LOW	MEAN	HIGH	LOW	MEAN	HIGH	LOW	N	N	N
1980 LEVEL*	$	$										
MULTIVERSITIES	3599.	5860.	1005.	7224.	8625.	5307.	3770.	7500.	1445.	11	4	11
UNIVERSITIES	2518.	5616.	925.	4318.	4650.	4000.	2673.	4520.	1425.	23	4	22
COLLEGE COMPLEXES	3025.	6250.	300.	6654.	9575.	2279.	2614.	6250.	300.	42	4	42
COLLEGES	3345.	6777.	925.	0.	4050.	4050.	0.	1717.	1717.	37	1	1
SPECIALIZED	3135.	5400.	847.	3933.	4280.	3520.	3923.	5400.	2580.	13	3	9
SEMINARIES	3029.	5150.	2000.	2168.	4650.	1015.	2054.	4170.	740.	6	11	11
HEALTH PROFESSIONAL	1386.	2197.	980.	6779.	9100.	3420.	2157.	4000.	1445.	3	4	5
NURSING SCHOOLS	2342.	3820.	1235.							6	0	0
TWO-YEAR SCHOOLS	1699.	4000.	670.							83	0	0
NEW YORK STATE	$2523	$6777.	$ 300	$4503.	$9575	$1015.	$2716.	$7500.	$300	225	31	106
1977-1979 % CHANGE**												
MULTIVERSITIES										0	0	0
UNIVERSITIES										0	0	0
COLLEGE COMPLEXES										0	0	0
COLLEGES										0	0	0
SPECIALIZED										0	0	0
SEMINARIES										0	0	0
HEALTH PROFESSIONAL										0	0	0
NURSING SCHOOLS										0	0	0
TWO-YEAR SCHOOLS										0	0	0
NEW YORK STATE										0	0	0

SOURCE: NYSED, POSTSECONDARY RESEARCH INFORMATION SYSTEMS 4/28/81

*1978 DATA USED IF 1979 NOT AVAILABLE
**1976-1978 DATA USED IF 1977-1979 NOT AVAILABLE

One of the problems with this approach is the volume of data involved. Profile charts may be used to summarize the differences between institutions in the system. Figure 2 and Table 5 show the percentage differences from the level mean and the simple differences from the percentage trend for each type of institution on each variable. As would be expected, universities and certain other types of institutions have higher salaries than do colleges, nursing schools and two-year schools. It may surprise some readers that multiversities and universities have less desirable percent tenured and faculty per program characteristics than do other types of institutions. It should be recognized that such a conclusion may indeed be false. It may be necessary to value a given percent tenured differently at universities than at colleges or two-year schools. Once appropriate value scales have been defined for each type of institution, then a figure such as this one would provide much more solid information. A great deal of supporting analysis and validation would be necessary to examine the assumptions on which both the value scales and the indicators themselves are founded so that usefulness of the procedures could be maximized.

The changes observed in variable characteristics between 1977 and 1979, as shown in the bottom half of Figure 2, are directly meaningful. For this example, a change greater than five percent was considered large. A change of less than one percent was considered to be no change. For salaries, percent changes were adjusted by 16 percent to account for inflation. Positive changes generally were present for the student per faculty and percent full-time characteristics of institutions. Interestingly enough, this may be the result of an economic downturn. As enrollments ease off and if cutbacks are initially made in part-time staff, then both of these

FIGURE 2: Selected Characteristics of New York State Institutions By Type of Institution

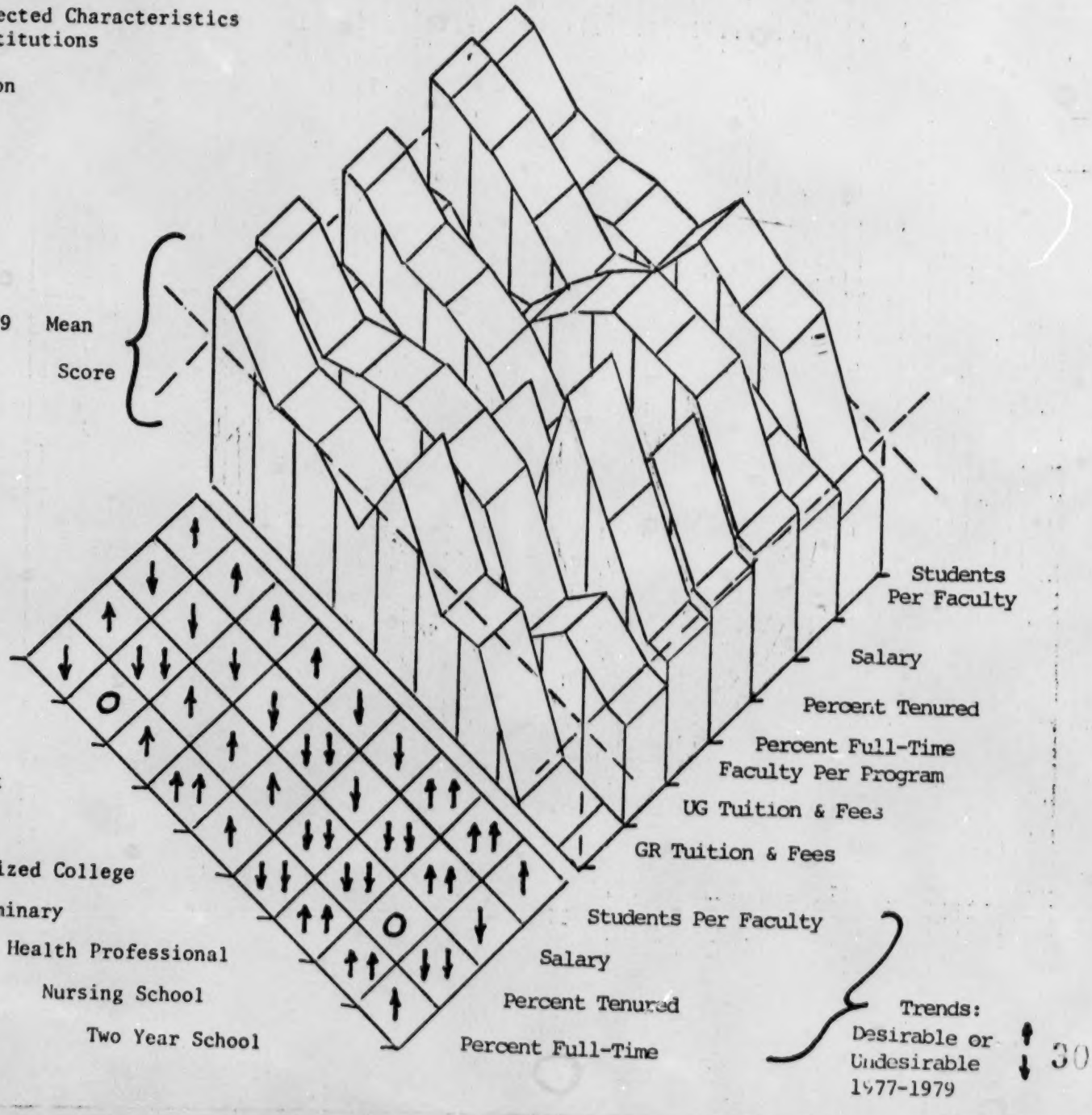


Table 5: New York State Comparison of Selected Institutional Characteristics
By Type of Institution[1]
1979 Level, Changes 1977-1979

	Students Per Faculty	Students/ Faculty With Ta's	Salary	Percent Tenured	Percent Full-Time	Faculty Per Program	UG Tuition & Fees	GR Tuition & Fees
New York State Mean	16.6%	16.1%	$22,615	40.6%	60.1%	8.2	$2,529	$2,777
Percent of Statewide Mean								
Multiversity	69%	57%	133%	140%	132%	65%	142%	136%
University	102%	98%	127%	163%	107%	55%	100%	96%
College Complex	99%	99%	112%	152%	109%	43%	120%	94%
College	91%	94%	90%	102%	101%	54%	132%	62%
Specialized	94%	88%	111%	59%	72%	118%	124%	141%
Seminary	43%	41%	83%	28%	87%	65%	120%	74%
Health Professional	64%	66%	105%	31%	88%	816%	55%	78%
Nursing	55%	57%	89%	0%	162%	140%	93%	-
Two-Year	128%	130%	90%	95%	95%	84%	67%	-
Change in Percent 1977 to 1979								
Multiversity	-3.4%	-5.6%	13%	-1.5%	-2.0%			
University	-4.2%	-5.4%	14%	7.3%	0			
College Complex	-3.2%	-2.4%	15%	1.8%	4.1%			
College	-1.9%	-1.6%	12%	1.4%	5.4%			
Specialized	-4.3%	-4.3%	7%	0.2%	-1.9%			
Seminary	-5.3%	-3.8%	19%	12.8%	-15.6%			
Health Professional	-8.5%	-6.2%	11%	12.4%	9.9%			
Nursing	-6.6%	-6.6%	24%	-	8.9%			
Two Year	-4.1%	-4.4%	12%	12.1%	2.4%			
New York State	-2.5%	-2.6	12.9%	5.6	1.6%			

[1] All institutional type measures are presented in terms of their ratio with the Statewide average for the measure for all types of institutions.

SOURCE: NYSED, Office of Postsecondary Research, Information Systems & Institutional Aid (GLR12/81)

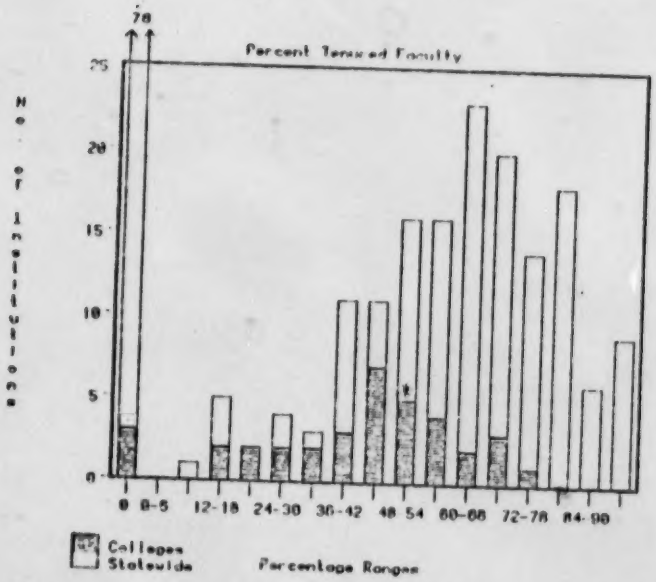
measures may change in desirable ways until the situation is severe enough to produce more than "trimming" actions by management. Negative changes are evident for both salaries and percent tenured levels. Only nursing schools experienced positive changes on the four variables measured. Negative changes on all variables were experienced on the average by specialized colleges and seminaries.

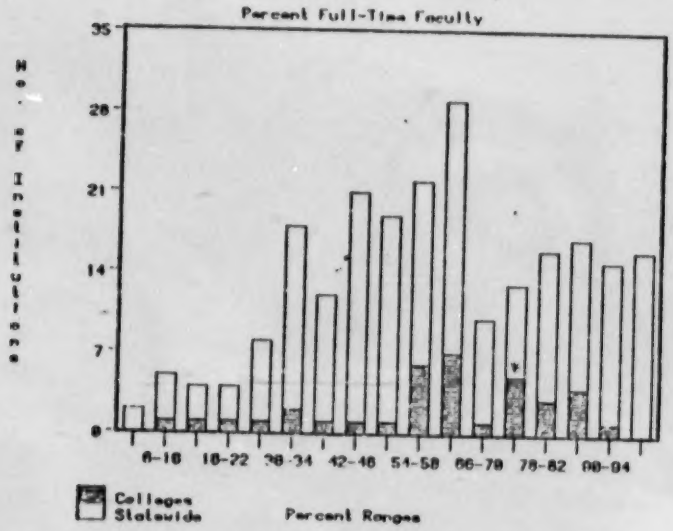
Other types of displays for examining systemwide characteristics on variables such as these are histogram charts and tables, and profile charts. Figures 3-5 on the next three pages illustrate the use of histograms and tables to display the distributions of institutional characteristics. Figure 3 shows how a relatively large set of specific faculty measures could be summarized on one page. That page would show the nature of distributions and how an institutional type distribution differs from the systemwide distribution. By the use of an asterisk, the position of a single institution can be referenced to the groups for each descriptor category. Standard deviations from the mean (z) or raw scores may be used on the X axis. This is necessary until more meaningful scales are defined. This raises again questions related to the assignment of value which will be discussed further in the next section.

An alternative means of highlighting the differences between types of institutions is provided by Figure 4. Histogram data for each type of institution on a single variable is presented next to each other. Differences in the distribution of each institutional type are readily evident and could serve as the basis for fruitful discussions by knowledgeable work groups. Computer programs have been developed that can automatically produce these tables.

Analysis of this sort can be extended further by developing profile graphs. These graphs illustrate the cohesiveness of groups of institutions across a

Figure 3


Percent Tenured Faculty
No. of Institutions
0 0-5 12-18 24-30 36-42 48 54 60-66 72-78 84-90
Colleges
Statewide
Percentage Ranges


Percent Full-Time Faculty
No. of Institutions
6-10 10-22 30-34 42-46 54-58 66-70 78-82 90-94
Colleges
Statewide
Percent Ranges

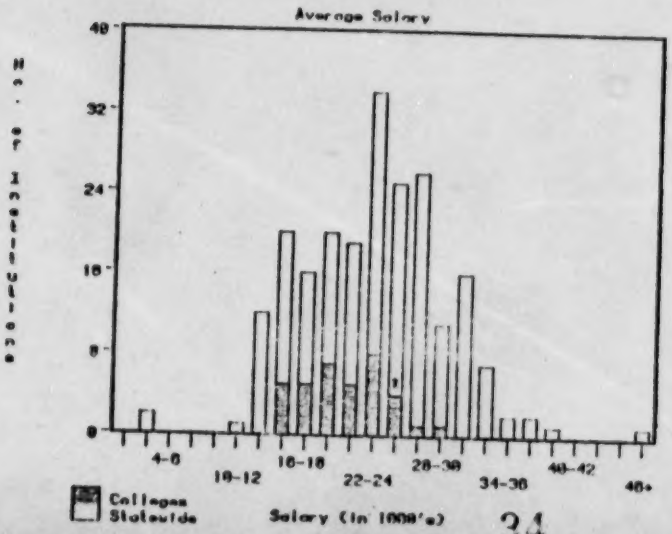

Average Salary
No. of Institutions
4-6 10-12 16-18 22-24 28-30 34-36 40-42 46+
Colleges
Statewide
Salary (in 1000's)

Figure 4

CELL COUNTS FOR PERCENT OF INSTRUCTIONAL FACULTY TENURED

Type of Institution

PERCENT TENURED	Multiversity	University	College Complex	College	Specialized	Seminary	Health Professional	Nursing	Two-Year	New York State
0%				2	9	9	2	6	34	62
5							1			1
10				1			1			2
15			1	3					1	5
20				2						2
25				1	2					3
30				2					1	3
35	1		3	3		1	1		1	10
40			2	4		1			0	7
45		2	4	5		1			1	13
50	3		3	6	1					13
55	1	(1)	5	2		1	1		3	14
60	1	5	6	2					6	20
65	2	3	2	3	1				4	15
70		5	5		1				2	13
75			6	1		1			8	16
80		1	3		3				8	15
85		1							3	4
90			1						7	8
95										0

(○) Denotes Institutional Position

set of variables. Figure 5 profiles universities in the State University, City University and Independent sectors. It is evident that universities within each sector are fairly cohesive and yet different from universities in other sectors. Independent universities are more diverse in nature than are SUNY or CUNY universities.

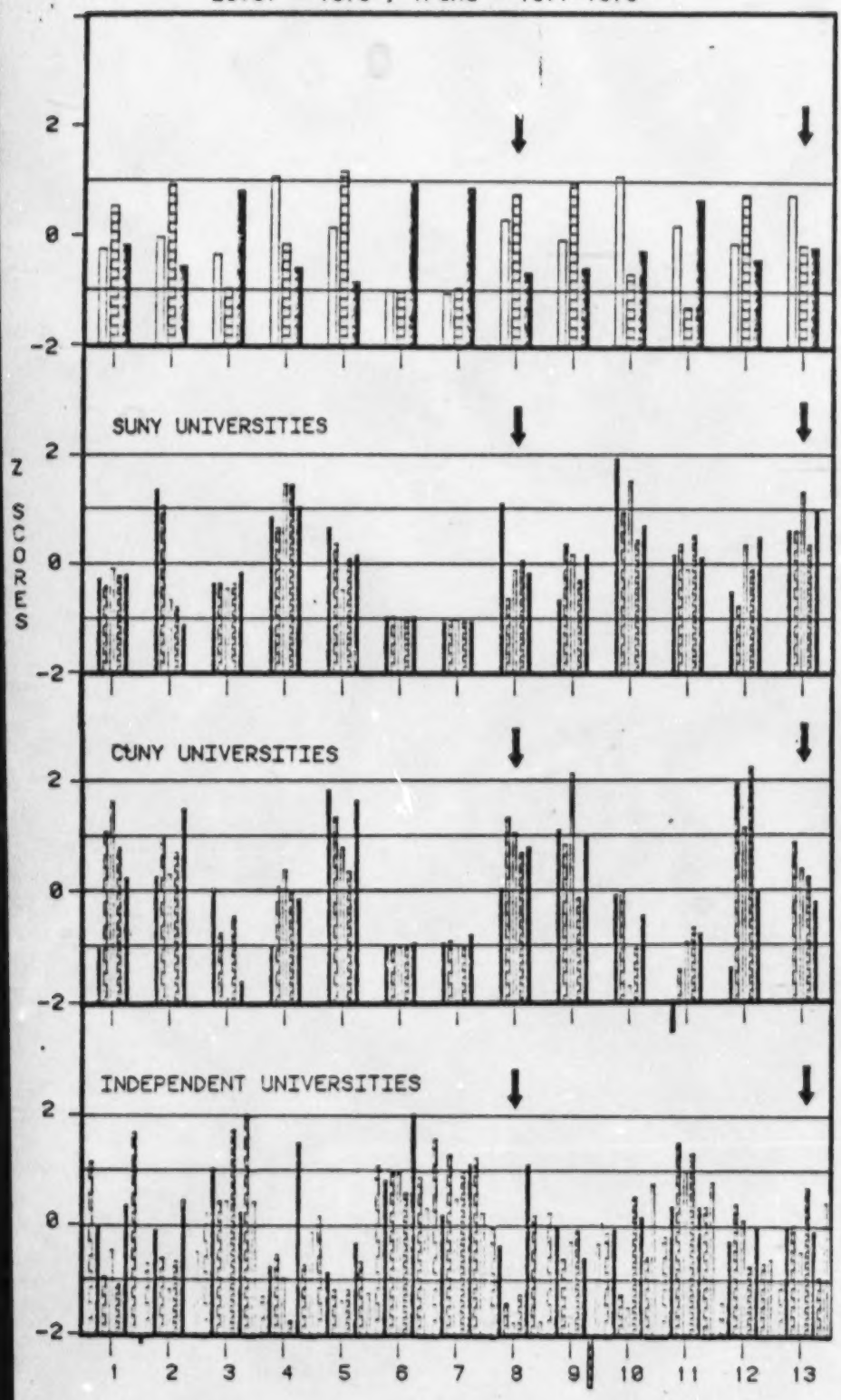
This type of information would be very useful for the development of typologies of homogeneous groups of institutions. The profiles also show that certain institutions are outliers on a few of the variables. These institutions deserve further analysis to expose problems in the institutions, or alternative operating arrangements that differ from conventional ones, data problems, or other unique circumstances.

Institutional Comparisons Through the Use of Z Scores: Comparison of institutional conditions across both institutions and variables is facilitated by converting all scores to a common scale. The most ideal scale, as stated earlier is an objectively grounded, interval quality scale. When a variable can be assumed to have monotonically increasing values then standard deviations from the mean score provide a useful scale. An institution near the positive extreme of the distribution on the majority of measured characteristics can be considered to be of relatively higher quality than most other institutions if the set of variables used is reasonably valid and comprehensive. Institutions near the negative extreme of the distribution on a substantial number of variables should probably be examined more closely. Special circumstances may be found which would overrule the conclusion of poor quality.

While a strong conclusion of either poor or high quality cannot be made from an analysis of standard deviations, the procedure appears to be highly effective for major discriminations. It can serve as an efficient screening device for identifying institutions which should be examined more closely. When it is

NEW YORK STATE
PROFILES FOR UNIVERSITIES By Sector And Institution
Level = 1979 ; Trend = 1977-1979

Figure 5



SUNY
CUNY
INDEPENDENT

SUNY ALBANY
SUNY BINGHAMTON
SUC BROCKPORT
SUC BUFFALO
SUC OSWEGO

BARUCH
BROOKLYN
CITY
HUNTER
QUEENS

ADELPHI
FORDHAM
HOFSTRA
LONG ISLAND
PACE
RPI
RIT
ST JOHNS

CURRENT LEVEL MEASURES

\- 1 Students/Faculty
\+ 2 Salary
\- 3 Percent Tenured
\+ 4 Percent Full-Time
\+ 5 Faculty/Program
\+ 6 UG Tuition & Fees
\+ 7 GR Tuition & Fees

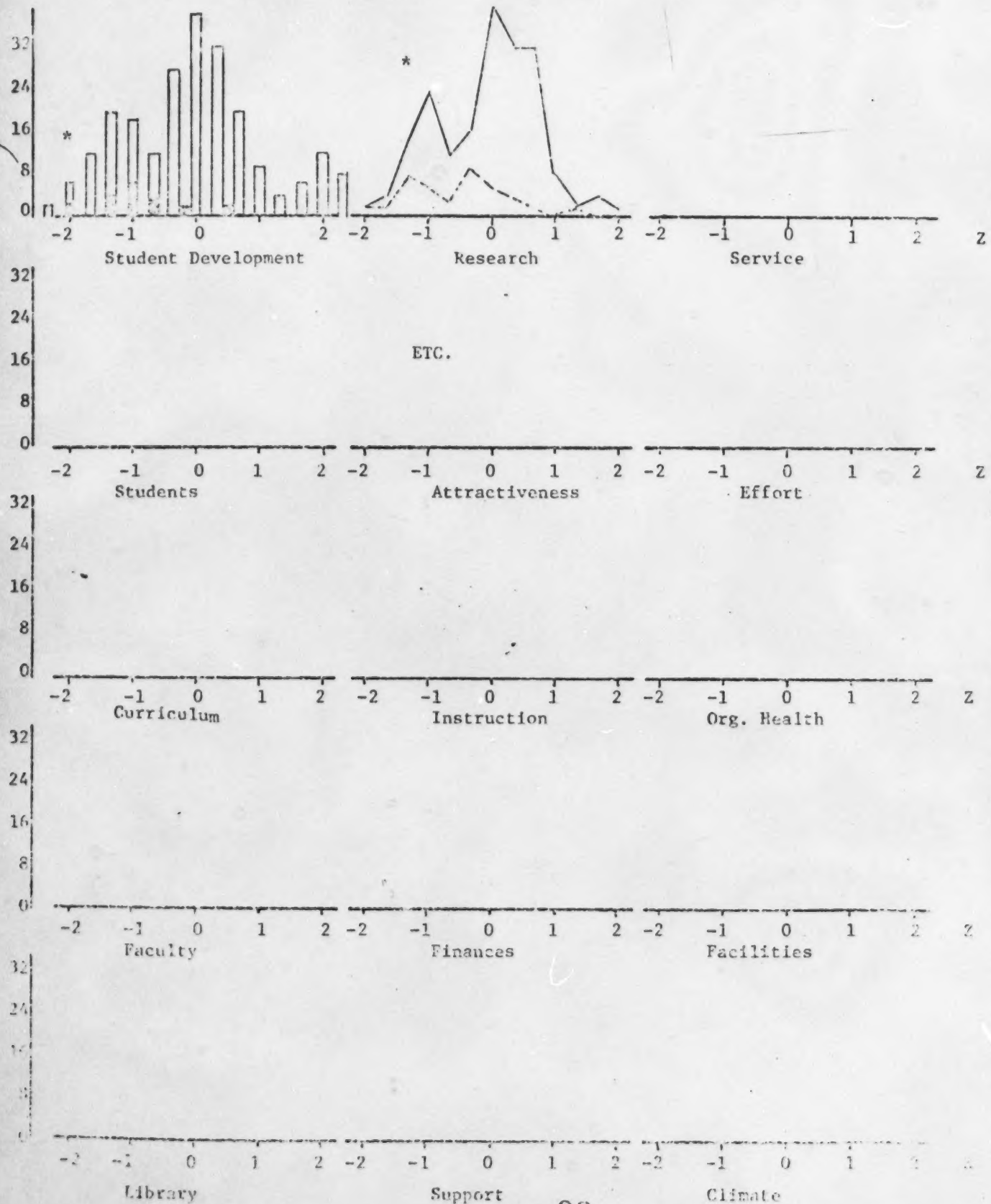
TREND MEASURES

9 Students/Faculty
10 Salary
11 Percent Tenured
12 Percent Full-Time
13 Mean Trend

reasonable to assume that scores on a variable will be distributed normally, one can expect that approximately 16 percent of the institutions will fall below one standard deviation from the mean and approximately 2.3 percent of the institutions will fall below 2 standard deviations from the mean. Under these conditions, standard deviation scores are typically referred to as z scores or standard units.

Research efforts related to this project have developed programs to calculate the z scores and also an average score across the full set of measures. Figure 6 suggests using hypothetical distributions, the kinds of broad profiles possible. The calculation of average scores is meaningful only if a utility measure with a direction of goodness can be assigned to each variable. Of course, an aggregate score should be developed in accordance with certain other principles typically espoused by scientific research methods. An aggregate score is generated to provide a complete and accurate summary. If our procedures were applied in a comprehensive and complete analysis, all the important factors that contribute to overall institutional quality (or other more specific outcomes) would be included as variables. Validity will be enhanced by measuring the factors accurately and by effectively distinguishing between factors and the measures of those factors. For instance, care must be taken to avoid including two measures of a factor when developing an aggregate score and thus double weighting the factor. For this reason, only of the two alternate student per faculty measures was one included when calculating the mean faculty resources score for the example used in this paper. In general, multiple factor measures should be aggregated in an earlier step. From this same perspective, it is important not to confuse <u>current</u> quality or conditions with a <u>trend</u> in quality or conditions. Level and trend measures should be aggregated separately.

Figure 6: STATEWIDE AND INSTITUTIONAL TYPE DISTRIBUTIONS OF Z SCORES FOR EACH DESCRIPTOR CATEGORY[1]



*denotes z score for this institution

[1] z based on statewide distribution

Distribution Statewide

Distribution for Universities

Figure 7 illustrates how one might summarize the position of an institution on the fifteen general variables identified in Figure 1. This hypothetical institution has particularly low scores on faculty resources, student characteristics, and attractiveness. The only above average characteristic which it has is its facilities, but changes in its facilities have been below average as well. Based on the mean level score, the institution is currently in the lower 16 percent of the institutions (mean level = -1.0z). Its recent trends have been highly negative (overall trend = -1.9z), in the lowest 2.5-3.0 percent of institutions. This institution is a prime candidate for further diagnostic analysis and possibly for visitation by an accrediting organization.

In this example, it should be noted that all z scores were calculated from the mean for the institution's own type of institution rather than the systemwide mean. Thus, the frame of reference is always comparable institutions and a given z score means the same across types of institutions permitting the development of rankings that include all types of institutions. However, statistical protocols require a preliminary step before ranking institutions. The calculation of a mean score across either descriptor categories (or measures within a category) changes the distribution properties of the score. Sixty-eight percent of the institutions will no longer fall between -1 and +1. The scores must be standardized again and must be submitted again as raw scores to the z score program before graphing the results on a z score scale.

The rankings for this example using faculty resources (done on New York State institutions) seem to conform well with conventional wisdom about the colleges. Furthermore, the rankings were generally consistent with groups of competing colleges as identified from an applications overlap study of student aid recipients (Rowse and Wing, 1981). Moreover, the z score ranks match well with competitive

Figure 7

INSTITUTION: _____

☑ Low on 2 or More Categories (Z Score <-2.0)

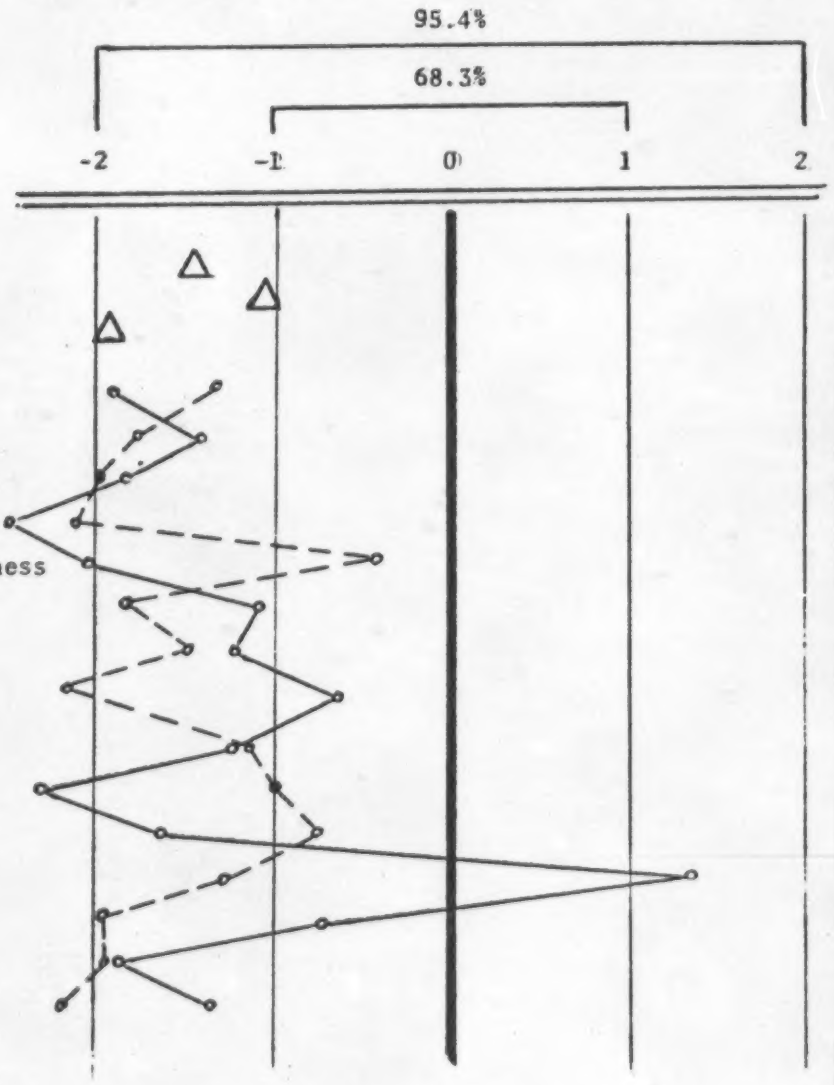
	Z Score
Faculty Resources	-2.30
Student Characteristics	-2.51
Attractiveness	-2.05

☐ Other: _____

SUMMARY OF SCORES[1] :

Proportion of Institutions between -Z+Z

Z Score :



Descriptor Category

Total
Level Total
Trend Total

1) Student Development
2) Research
3) Service
4) Student Characteristics
5) General Environment: Attractiveness
6) General Environment: Effort
7) Curriculum
8) Instruction
9) Organizational Health
10) Faculty Resources
11) Financial Resources
12) Facility Resources
13) Library Resources
14) Support Structures
15) Institutional Climate

[1] Z based on institution type distribution

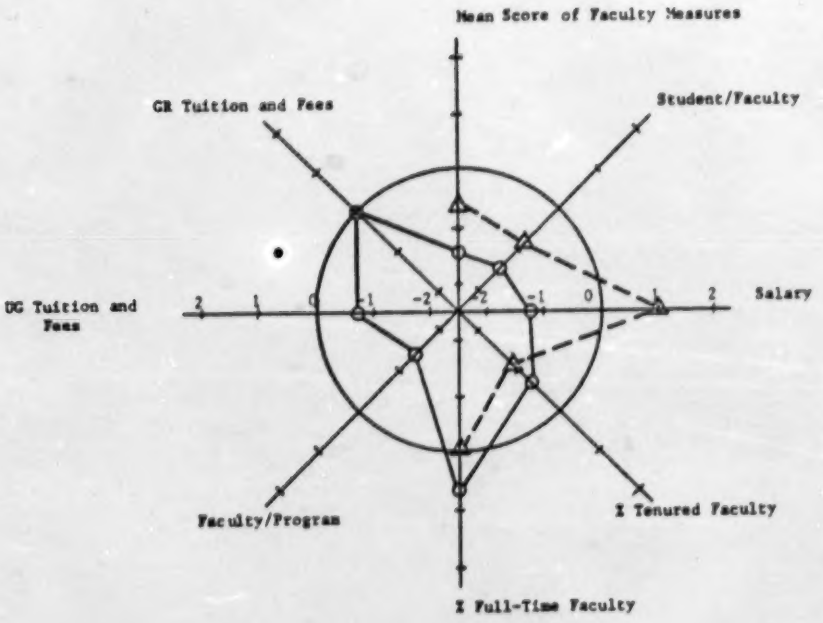
dominance ranks within each group as defined by acceptance ratios, draw ratios and mean SAT scores. It appears that some very useful explanatory information can be achieved even when using only a few variables (i.e., faculty and salary). It is reasonable to expect that a more complete set of variables would provide some very reliable rankings. It remains to be seen which set of variables will produce rankings that correlate best with quality, attractiveness, or some other outcome.

A more detailed analysis of an institution would involve looking at each of the measures included in each descriptor category. The information could be summarized in a profile similar to the one in Figure 7. An alternative presentation is shown by Figure 8. This shows results on seven of the eight measures included in this example for a low ranking and a high ranking institution using actual data from New York State. A variation of polar axes is used to facilitate visual analysis of the data. Each axis is oriented so that an average score on the variable (z=o) lies on the circle. Tic marks represent a change of 1 z score. Lower than average scores fall inside the circle and higher than average scores fall outside of the circle. Thus a better quality institution should produce a fairly large polygon and a poorer quality institution should produce a more compact polygon. The area of the polygon might even be considered as an aggregate measure. As before, a separate polygon is developed for both the current condition and the trend. In this case, trend measures were not available for the faculty per program and the tuition variables. It is probably typical that an institution with high current levels has lower trend scores and vice versa. This is an example of the regression to the mean effect.

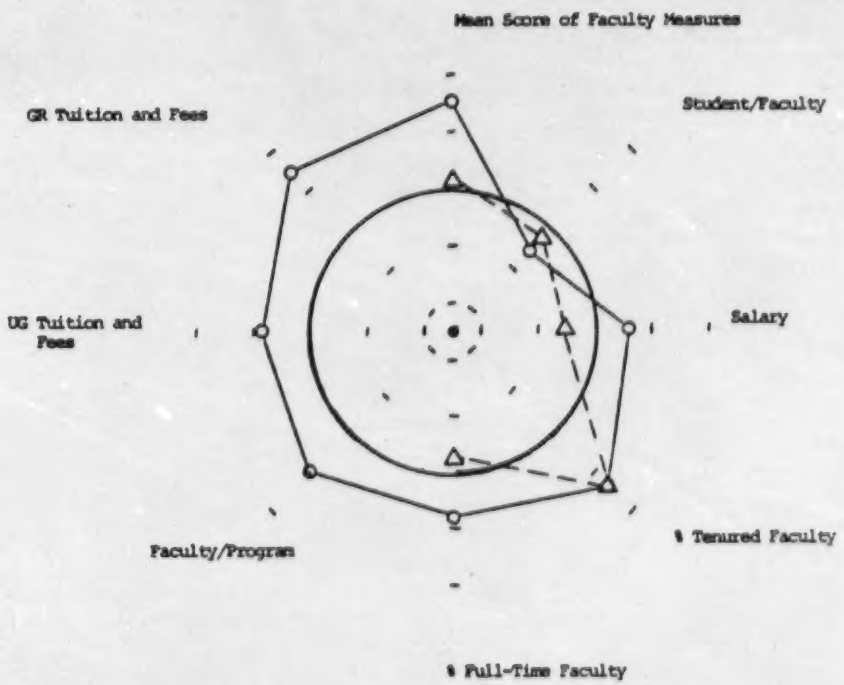
Figure 8

Z SCORES WITHIN INSTITUTIONAL TYPE

A LOW RANKING INSTITUTION



A HIGH RANKING INSTITUTION



Correlational Studies: Both the development of measures to be included in a descriptor category (e.g. faculty resources) and the validation of a relationship between a descriptor category and some specific outcome can benefit from correlation analysis. Correlations were calculated between all of the measures included in this example. Both raw score and z score measures based on institutional type were included to illustrate the differences associated with the alternate measurement forms. Figure 9 summarizes the correlation matrix. The actual correlations are shown in Table 6.

In general, z scores produced higher correlations between a variable and the mean faculty score than did raw scores. The differences were not large but were present for both the trend and level measurements. The average z based correlation between a trend variable and the mean faculty trend score was .418. Using raw scores the correlation was .372. For the current level measurement, the respective correlations were .326 and .258. These results indicate that analyzing the measures within the institutional type defined for this study (z score based on institutional type) is preferable to working with raw scores across institutional types. Thus, the remainder of the discussion concentrates on the correlation between z score measures (the lower right portion of the matrix in Figure 8).

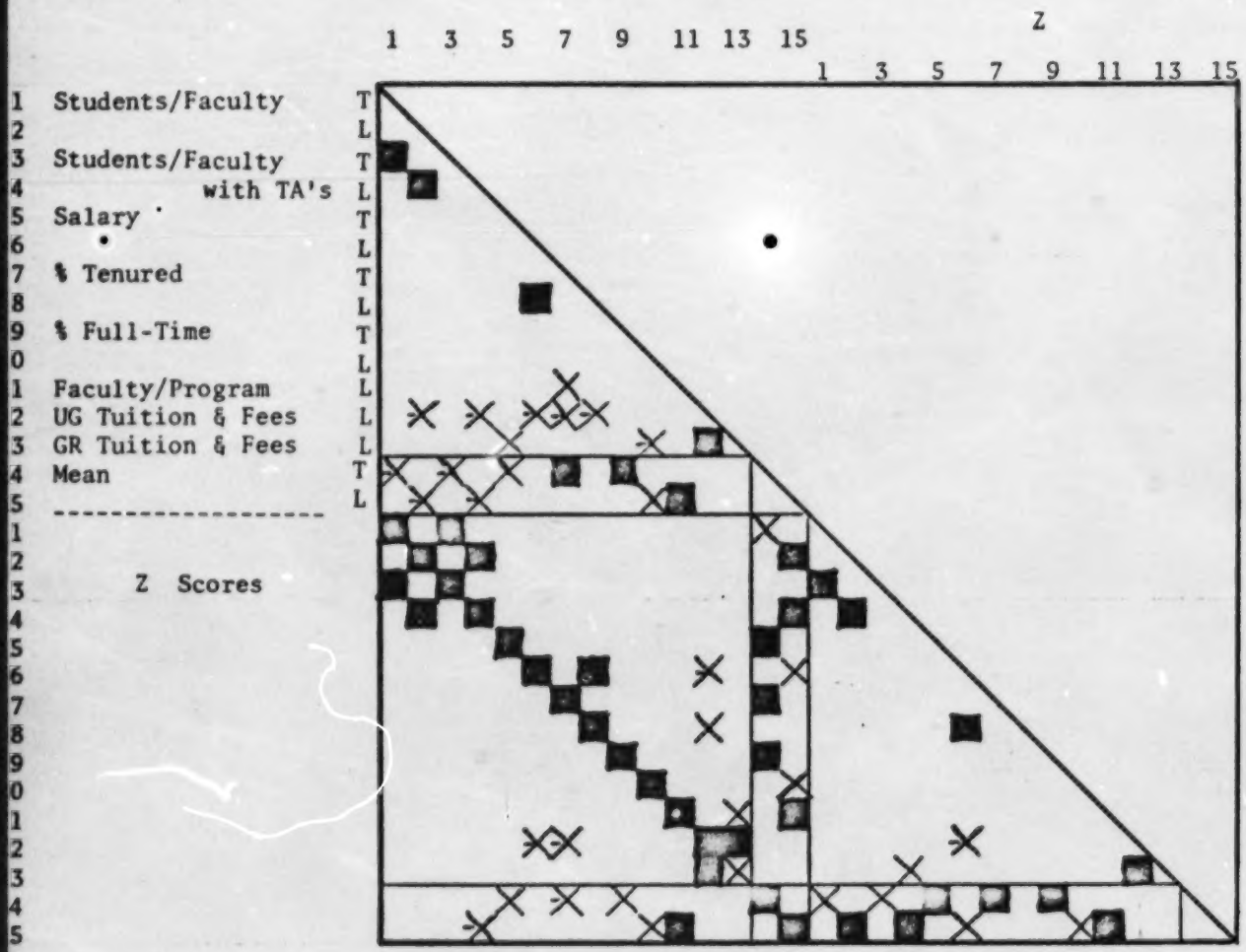
A second conclusion that may be drawn from this matrix is that, at least for these variables, the trend and current level measures are distinct factors. The absolute value of correlations between trend and level measures on the two bottom rows of the matrix average only 0.09. It is important then to develop rankings of institutions separately for the current level and trend measures.

As stated earlier, the average correlation with the overall faculty score was .418 and .326 for trend and current level, respectively. These are moderate correlations at best. Since faculty resources is not a simple attribute but rather

FIGURE 9

FACULTY MEASURES

CORRELATION MATRIX



□ Correlation less than .30

⊠ Correlation Between .30 & .40

■ Correlation Greater than .40

T: Trend = % Change 1977-1979

L: Level = 1979 data, 1978 if 1979 missing.

Table 6 CORRELATION MATRIX
(Faculty and Tuition Measures)

1	Students/Faculty		T													
2	Students/Faculty		L	.26												
3	Students/Faculty with TA's		T	.96	.22											
4	Students/Faculty with TA's		L	.20	.97	.19										
5	Salary		T	.12	-.03	.11	-.03									
6	Salary		L	.01	.02	-.02	-.05	.25								
7	% Tenured		T	-.18	-.03	-.14	.00	.18	.06							
8	% Tenured		L	.16	.24	.13	.24	.18	.43	.14						
9	% Full-Time		T	.04	.04	.10	.07	-.06	-.10	-.03	-.21					
10	% Full-Time		L	.18	.03	.22	.06	.15	.02	-.08	.07	.24				
11	Faculty/Program		T	-.05	-.07	-.02	-.06	.02	.15	.30	-.11	-.06	-.02			
12	UG Tuition & Fees		L	-.10	-.33	-.01	-.32	-.07	-.31	-.30	-.33	-.04	-.10	-.06		
13	GR Tuition & Fees		T	.11	-.26	.01	-.28	.30	-.10	-.13	.00	-.20	-.32	-.27	.88	
14	Mean Score		L	-.30	.00	-.30	-.01	.40	.06	.43	-.12	.43	.01	-.07	.03	-.03
15	Mean Score	Value Assigned	T	-.07	-.30	-.03	-.31	.01	.24	.12	-.06	.04	.31	.58	-.09	-.17
1		-	L	-.85	-.20	-.84	-.19	-.09	.03	.15	-.11	-.14	-.21	.04	.05	-.05
2		-	T	-.23	-.73	-.21	-.70	-.09	-.04	.10	.10	-.16	-.08	.07	.16	.22
3		-	L	-.81	-.18	-.88	-.19	-.08	.03	.13	-.10	-.16	-.24	.01	.04	-.02
4		-	T	-.20	-.70	-.21	-.71	-.08	-.02	.07	.09	-.17	-.09	.06	.16	.23
5		+	L	.10	.06	.08	.05	.82	.20	.16	.17	-.10	-.01	.01	-.08	.03
6		+	T	-.03	.05	-.06	.01	.17	.76	.17	.44	-.12	.00	.18	-.30	.04
7		-	L	.15	.08	.13	.05	-.12	-.08	-.91	-.13	-.01	.05	-.26	.24	.14
8		-	T	-.01	-.10	.03	-.09	-.14	-.27	-.18	-.71	.21	-.08	.01	.30	-.06
9	Z Scores	+	L	.09	.07	.12	.08	-.07	-.07	-.05	-.12	.88	.17	.00	-.09	-.14
10		+	T	.22	.11	.23	.07	.08	.07	-.09	.10	.19	.83	.00	-.20	-.24
11		+	L	-.05	-.04	-.02	-.03	.01	.16	.24	-.03	-.02	-.02	.81	-.18	-.34
12		+	T	-.13	-.09	-.12	-.08	-.11	-.33	-.30	-.28	-.07	-.02	.07	.58	.53
13		+	L	.14	-.10	.19	-.10	.08	-.06	.01	-.11	-.13	-.17	-.15	.47	.36
14			T	-.28	-.04	-.28	-.05	.39	.08	-.38	-.12	.39	-.00	-.05	.03	-.09
15			L	-.07	-.29	-.04	-.30	.00	.24	.11	-.07	.05	.31	.58	-.09	-.18

2		.21														
3		.94	.19													
4		.19	.97	.20												
5		-.08	-.11	-.04	-.07											
6		.09	.00	.10	.03	.21										
7		-.14	-.10	-.15	-.06	-.10	-.10									
8		-.01	-.10	-.03	-.09	-.17	-.50	.22								
9		-.14	-.18	-.17	-.19	-.14	-.09	.04	.10							
10		-.21	-.12	-.21	-.08	.02	.06	.09	-.20	.21						
11		.10	.08	.05	.08	-.03	.21	-.29	-.05	.07	.01					
12		.04	.15	.07	.16	-.15	-.34	.27	.25	-.14	-.08	-.08				
13		-.11	.29	-.15	.31	-.07	.05	-.06	.21	-.18	-.17	-.01	.83			
14		.36	-.10	.32	-.07	.46	.08	.49	.09	.46	.06	-.03	-.06	-.27		
15		.10	.43	.06	.45	-.03	.33	-.10	.14	.08	.37	.70	.02	.17	.06	

a complex aggregate of not necessarily independent contributors, this result is not bothersome. The items contributing to the mean score are essentially uncorrelated as indicated by the large area of correlation less than .30 in Table 6. This is a desirable characteristic. One exception is the correlation between students per faculty and students per faculty including teaching assistants (r = .94 and .97). These two measures should be aggregated or a selection made between them (as was done in this analysis) before entering them into the total faculty score. Another exception is the correlation of +.50 between salary and percent tenured. We may be confusing two sub-factors by our choice of variables included. Percent tenured is intended to reflect the structure of faculty types and consequently the flexibility possible in meeting changing faculty and financial needs. Salary was intended to reflect the resources available to attract top people and therefore the likely types of faculty present. However, since tenured faculty have above average salaries, the higher the percent tenured, the higher the expected salaries. A better measure than salary which does not likely correlate with percent tenured may be salary of faculty at a given academic rank. Perhaps this could be explored in future work.

Moderate to low correlations also occurred for salaries and undergraduate tuition (N = -.34) and students per faculty with TA's and graduate tuition (r = -.31). It is unclear what would produce the first correlation, but the second may flow from the research and graduate emphasis of institutions and the typical consequences for teaching emphasis. Undergraduate and graduate tuition are also highly correlated but otherwise uncorrelated with the faculty measures. Faculty resources and tuition charges reflect the two major factors included in this analysis. Measures for one should not be included in the total score for the other.

These examples suggest the general utility of examining the correlations between variables. Several extensions of the analysis would be appropriate. First of all, a more comprehensive set of variables would provide a much clearer picture of the higher education system. Second, interaction variables should be constructed and examined for importance. It can be easily hypothesized that certain combinations of variables have unique influence that would be unpredicted from the individual variables. Third, and most important, hard outcome measures for higher education need to be collected and included in this analysis. Only in this way can we be sure that our assumptions about the influence of variables on quality are indeed valid and appropriate.

POTENTIAL BENEFITS OF PROPOSED PROCEDURES

The assessment procedures described above have potential benefits for a number of groups including individual campuses, state educational agencies, as well as the higher education community in general. The benefits from each of these perspectives are presented below.

Benefits for Institutions

1) Most institutions are concerned about quality and performance, both for professional reasons and for reasons of institutional survival. Most institutions would welcome any approach that helps them make better academic planning decisions as long as it does not jeopardize their position. The proposed procedures seem to offer clear prospects for better academic planning, although there may be uncertainty about their impact on survival and competition.

2) The project will supply institutions with the tools for assessing quality and the means to utilize those tools in academic planning. Currently most institutions lack the resources to develop these tools on their own. From an institutional point of view, this framework enables three kinds of comparisons: 1) historical/longitudinal self-comparison on the basis of trend data; 2) "snapshot" comparison with cutoffs or ranges which have been established as normative; and 3) comparisons with a group of other similar institutions.

3) Through the use of common indicators, data collection procedures, and presentations, institutions can obtain more comparative data than is currently available. The proposed procedures would also force a mission/goal orientation on quantitative assessment that could make it attractive.

4) The institutions recognize that if they are not more proactive in planning for decline, other groups may usurp their prerogatives.

Benefits for State Agencies or Superordinate Coordinating Bodies (Multi-campus systems and Consortia)

1) Some state agencies are directly involved in program review and others would like better information about quality and outcomes for budgeting processes. The results of these procedures would support these functions, particularly in the initial screening stages.

2) State agencies currently have little basis for making judgments about quality. A set of indicators which have been validated and can be used to improve assessments of institutional quality or performance would be extremely helpful in that process.

3) Data based indicators can be used to support policy decisions and statewide planning as well as program or institutional assessment.

4) This project provides a framework for making planning decisions on a rational and systematic basis rather than a political one.

Benefits from the Perspective of the Entire Higher Education Community

1) The procedures, even if controversial, will act as a catalyst to focus attention on the need for more effective and widely accepted methods for assessing quality. Few new ideas have been developed in recent years.

2) The higher education community will increasingly be required to justify itself to society both in terms of quality and impact. The proposed procedures can provide a basis on which greater accountability can be achieved and better justifications for resources can be developed.

3) The management of contraction cannot be accomplished solely on the basis of institutional self-interest. The higher education community must demonstrate the capacity for cooperative planning for contraction or that process will be preempted by other interest groups. The proposed procedures can supply the tools to implement cooperative planning.

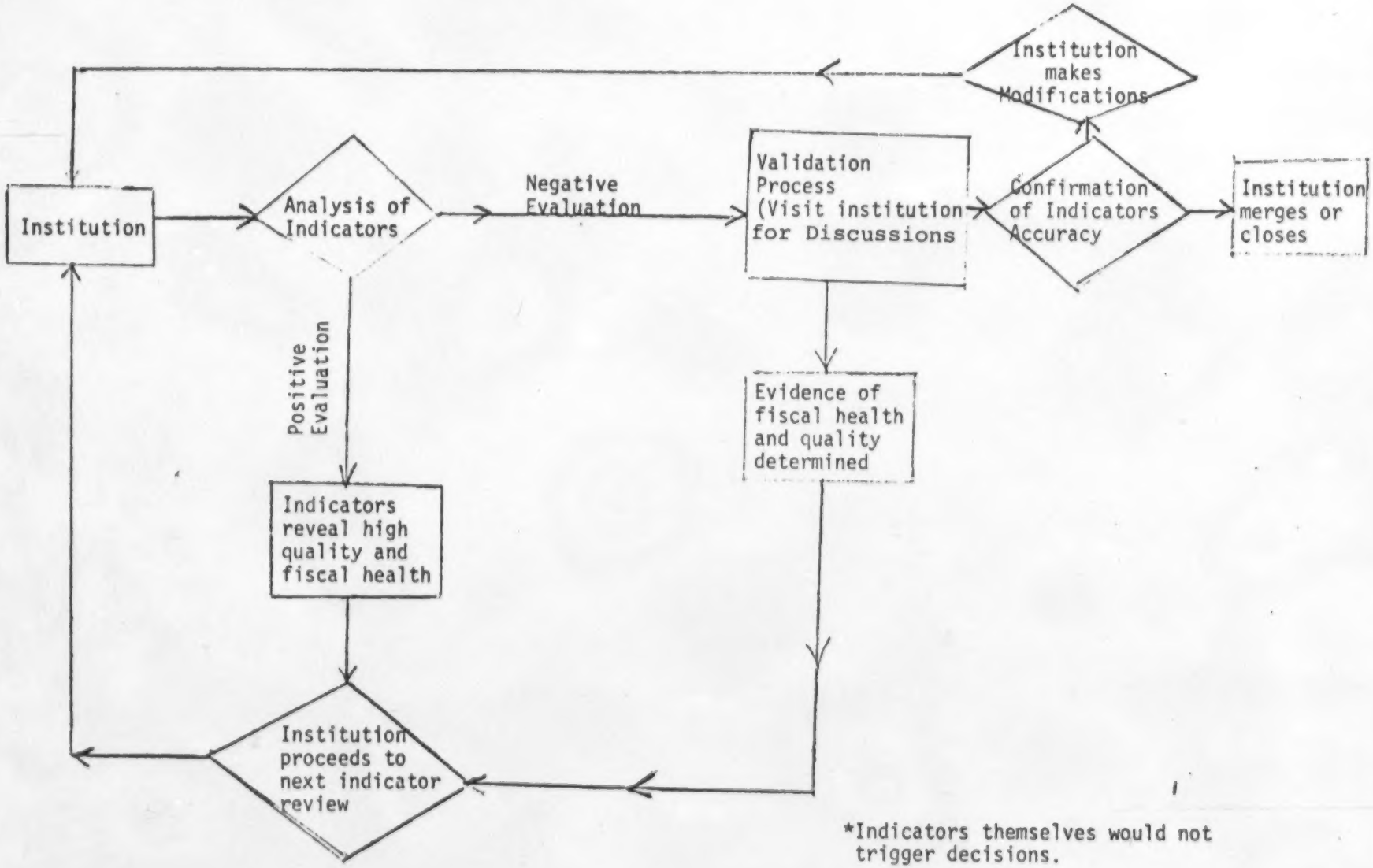
THE USE OF THE QUALITY ASSESSMENT PROCEDURES IN INSTITUTIONAL SCREENING

One of the key tests of the value of the procedures described in this report is applicability in decision making at either a statewide planning level or within an institution. Figure 10 on the following page provides a visual representation of how the approach described in this paper might be used in an institutional evaluation process.

At the first step in the process, the institution reviews the indicators in their entirety. If this review reveals that the indicators in the aggregate (a z score derived from a combination of all the available indicators) or a significant number of specific indicators deviate from the norm in a negative direction, then it may be presumed that deficiencies may exist. It may also be evident that problems are manifest because trend data suggest a decline on key variables relative to other institutions, even though the current levels on these variables are reasonable. Consequently, even though the institution may be within acceptable tolerances at the time of the analysis, its comparative advantage may be declining. If the review is positive, obviously, no further action is necessary (beyond those activities that sustain the institution's vitality and quality).

If this initial statistical review indicates a negative position, then a more extensive review will be necessary to determine if there are in fact serious problems at the institution, and if so, precisely what the problems are. The quantitative indicators in this respect can be a valuable diagnostic tool. They can identify the specific areas that may need attention and remedy. It should be noted that in Figure 10 this activity is referred to as a "validation process" because a comprehensive review will either confirm that problems exist and therefore validate the utility and accuracy of the indicators in highlighting problems, or it may reveal that the indicators incorrectly identified problems. If the

Figure 10 Screening Model Based on Quality Indicators Project


Institution makes Modifications
Institution
Analysis of Indicators
Negative Evaluation
Validation Process (Visit institution for Discussions
Confirmation of Indicators Accuracy
Institution merges or closes
Positive Evaluation
Evidence of fiscal health and quality determined
Indicators reveal high quality and fiscal health
Institution proceeds to next indicator review

*Indicators themselves would not trigger decisions.

indicators are not confirmed and it is concluded that no serious problems exist, then the institution proceeds on its projected course. If the validation process confirms the accuracy of the indicators in identifying problems, then the institution must make the modifications necessary to remedy deficiencies before it proceeds. In either case feedback from the validation will be used to adjust the procedures, measures, and cut-off values as appropriate.

This process can be undertaken entirely at the institutional level. It can also be used as an evaluation device by statewide regulatory agencies or voluntary accrediting organizations. In this instance, what can be envisioned as a "layered screening process" may be utilized. At the first layer or level of review, an aggregate analysis is done. Those institutions falling below a pre-specified aggregate score across all variables and measures would be selected for further review. Key variables which correlate highly with other judgments about quality could then be scrutinized. (Further research is necessary to identify which variables do in fact correlate most highly with severe institutional deficiencies). If these variables demonstrate that the institution significantly exceeds norms for its category, then this should trigger a comprehensive on-campus evaluation. This evaluation would rely on expert judgment to make decisions related to the actual existence of low quality and can suggest specific remdiation efforts. It should be clearly understood that a review of the indicators would never trigger anything other than closer scrutiny. They should never be utilized as a substitute for expert judgement, and should not be used as the sole basis on which critical decisions for an institution are made.

ADDITIONAL WORK FOR THE FUTURE

A number of tasks remain to be addressed. They represent a possible agenda to which future resources could be directed.

1) Increase the number of measures (variables and descriptors) used in the matrix.
2) Enlarge the data base by collecting data on categories and descriptors for which data are not currently available.
3) Begin collecting longitudinal data for long term validation on such things as student self-assessment and value-added.
4) Examine correlations between the proposed framework and other measures of quality (e.g., expert judgements) to connect the indicators project to "real world" assessments of quality.
5) Employment of panels of experts to review project.
6) Develop strategies by which planning activities and policy analysis can effectively utilize the model.

There are several features about this project that are unique and warrant further consideration. First, as has been pointed out above, this project, unlike other efforts to assess quality which appear in the literature, is multi-dimensional in that it examines not only student development (the area which outcome studies typically try to address) but "research" and "service" as well. Ultimately, the full range of possible programs and missions should be dealt with. Furthermore, because of the use of a systems approach, it has not been necessary to divorce outputs from inputs and process variables. Given that unequivocal or routine measurement of outcomes and quality will probably never be possible, it is important not to discard inputs and process variables from our measurement tool box. In some cases these are the only available means for assessing quality. But, we expect them like all proxy measures, to reveal a high positive correlation with other measures focusing on the same issue; hence the need for additional work on longitudinal data.

Finally, as has been demonstrated in the illustrations using faculty resource statistics, the project will permit comparisons based on historical longitudinal self-comparison from trend data, snapshot comparisons with cutoffs or ranges which have been established as normative and comparisons with a group of other similar institutions. These aspects of the procedures are unique and would be a significant advantage in the formulation of policy both at the institutional and regional level.

SUMMARY AND CONCLUSIONS AND CAVEATS

This report describes procedures that might be used to incorporate quantitative information into higher education assessment processes at the campus or state levels. While the report suggests that quantitative indicators and measures may have a very useful place in evaluation and assessment processes, particularly in screening or trigger mechanisms, it should not be construed as an endorsement of such meausres as the sole basis for important decisions.

In addition to suggesting a comprehensive framework and procedures for incorporating quantitative information into assessment processes, the report provides examples of indicators, measures, statistical displays, and interpretations of results that clearly describe the proposed approach. Specific illustrations are provided for variables and measures related to faculty resources using actual data from colleges and universities in New York State.

Caveats

Some final observations are necessary to underscore the limitations of the model of quality assessment that has been described here:

1) This paper reflects tentative and incomplete efforts of a research team whose work is still in progress. All claims for the utility of the model are offered with great caution and with the intention that they will be shared primarily with other researchers in this field. The measures and procedures should be validated prior to introduction into actual assessment processes.

2) The higher education community has exhibited strong reservations about the use of quantitative data in quality assessment in the past. These reservations derive in large measure from the reductionism that quantitative analysis so often reveals. This model, however, is intended only as a supplement to the broader philosophical aspects of traditional quality assessment as practiced in the field today.

3) Any effort to rank the institutions in a state or region by using aggregate scores derived from this model would be inappropriate and irresponsible. Because there is such uncertain correlation between the variable, a composite score as an overall quality rank would be meaningless without prior validation of the measures and procedures.

4) Although a central agency like the New York State Education Department may well be appropriate as a center for data collection and dissemination, it should not on the basis of that role, be construed as an arbiter of quality. The proposed procedures are intended primarily as diagnostic tools to aid state planners in identifying problems that must be dealt with.

References

Gourman, Jack. The Gourman Report. Los Angeles: National Education Standards, 1977.

Hester, Donald G. Postsecondary Education - Towards a Definition of Quality. Background Paper. Albany: OPRISIA, State Education Department 1981.

Rowse, Glenwood L., and Wing, Paul. "Assessing Competitive Structures in Higher Education." Journal of Higher Education forthcoming.

APPENDIX A

Example

Measures of College Status & Performance & Cutoff Criteria For Review Assessments

The two colleges described on the following pages represent two possible types of colleges. They are hypothetical examples. Measures are suggested for each of the variables presented in Figure 1. A criterion or trigger for a screening device for campus visitations and review are suggested for each measure. These triggers are set specifically for a given college type and mission or purpose category. Given limited knowledge, it was frequently necessary to specify the general rather than specific nature of a trigger (e.g., a low value rather than less than 20 percent). One measure may be appropriate for more than one purpose and would have a distinct triggering level for each purpose.

The three utilization columns suggest the emphasis which different groups might choose to give to each measure. Campuses tend to be more interested in formative evaluation and thus might delve into more specific process and outcome measures than would other groups. State agencies might take a variety of approaches. They tend to focus on minimum standards with some attention to excellence. Columns 1 and 2 for state agencies simply represent examples of different levels of commitment to using quantitative measures and triggers.

APPENDIX B

Definitions of Indicator Variables

1) STUDENT/FACULTY RATIO

   a) Level and Trend of FTE STUDENTS/FTE INSTRUCTIONAL FACULTY. FTE students were calculated using FTUG _ (1/3 * PTUG) + (.4 x PTGR). FTE Instructional Faculty are Full-Time Instructional Faculty + (1/3 x Part-Time Instructional Faculty). Trends are a simple percent change from first to third year of a three year range.

   b) Same as a) except FTE teaching assistants are included in the FTE Faculty measure.

2) FACULTY/PROGRAM RATIO

   a) LEVEL of all programs leading to a certificate or degree divided by FTE Instructional Faculty plus teaching assistants. FTE faculty are calculated the same way for all variables.

3) AVERAGE SALARY

   a) 10 month contracts were annualized by a factor of 1.2, then added to 12 month contracts and divided by Full-Time Instructional Faculty (All data in IF10-IF12 records). Level and Trend. (Trend = % change).

4) PERCENT FACULTY WITH TENURE

   a) Level and Trend of Tenured Faculty divided by Total Full-Time Instructional Faculty. All data for this ratio is from IF10 and IF12 records, (Trend = % Change).

5) PERCENT OF FACULTY THAT ARE FULL-TIME

   a) Level and Trend of Full-Time Instructional Faculty divided by total Instructional Faculty. All data is from APFT and APPT records.